DOCUMENT RESUME

ED 093 973 TM 003 828

AUTHOR Bryk, Anthony S.
TITLE The Brookline Early Education Project: Resolving Methodological Issues in Evaluating an Early Childhood Education Program Model.
PUB DATE 73
NOTE 18p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, Illinois, April, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Cost Effectiveness; Early Childhood Education; *Educational Diagnosis; *Evaluation Methods; Identification; *Models; *Preschool Programs; Program Evaluation; *Research Methodology
IDENTIFIERS BEEP; *Brookline Early Education Project

ABSTRACT

The Brookline Early Education Project (BEEP) was initiated to determine the cost effectiveness of providing comprehensive diagnostic and educational services for 285 very young children and their parents. The primary goal of the program is to bring the family, the school, and the medical profession into a working relationship early in the child's life, thus hoping to ensure maximal opportunity for success in the school years. In planning BEEP, a number of methodological issures arose. This paper discusses the resolution of the following three issues: (1) the control group, (2) the generalizability of results, and (3) the relative importance of process versus outcome assessment. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 093973

TM 003 828

THE BROOKLINE EARLY EDUCATION PROJECT:

Resolving Methodological Issues In Evaluating An Early Childhood Education Program Model

(ABSTRACT)

The Brookline Early Education Project (BEEP) was initiated to determine the cost effectiveness of providing comprehensive diagnostic and educational services for 285 very young children and their parents. The primary goal of the program is to bring the family, the school, and the medical profession into a working relationship early in the child's life, thus hoping to ensure maximal opportunity for success in the school years.

In planning BEEP, a number of methodological issues arose. This paper will discuss the resolution of the following three issues: (1) the control group, (2) the generalizability of results, and (3) the relative importance of process versus outcome assessment.

ED 093973

THE BROOKLINE EARLY EDUCATION PROJECT: Resolving Methodological Issues In Evaluating An Early Childhood Education Program Model

Presented at the 1974 Annual Meeting of the National Council on Measurement in Education, Chicago, Illinois: April, 1974

Anthony S. Bryk
Brookline Early Eduaation Project
40 Centre Street
Brookline, Mass. 02146

A number of methodological issues arose in developing an evaluation design for the Brookline Early Education Project. This paper discusses our resolution, of the following three concerns: (1) the control group, (2) the generalizability of results, and (3) the relative importance of process versus outcome assessment. These issues are not new; they have constantly plagued researchers in this attempts to evaluate the impact of various social innovations. In order to better understand our approach to these questions, we present first a brief overview of the project - its history, and the nature of the services we provide for parents and their children. In this light, we then examine the methodological issues raised above.

BACKGROUND

Public schools across the country are now concerned about developing programs which ensure that children have an optimal chance for success in school and later life. In Brookline, Massachusetts, the public schools have developed a model which could have nationwide implications for early education and for reduction of educational handicaps.

The program, entitled the Brookline Early Education Project and dubbed BEEP, was initiated by the Superintendent of Schools, Dr. Robert I. Sperber. He was concerned that public education spent more funds on the education of high school students than on the education of kindergarten and primary age children, despite the evidence that cognitive development was fairly well established by the time youngsters were eight years of age. He was also con-

cerned about children who enter school with a learning handicap or inefficiency and who, despite an exceptionally capable teaching staff and the best available remedial programs, become caught in a cycle of school failure.

Dr. Sperber consulted Burton L. White, Director of the Preschool Project at the Harvard Graduate School of Education. Over the past decade, the Preschool Project had studied the developing competencies of many young children and the conclusions (White and Watts, 1973) certainly had pertinent implications for Brookline. In particular, it seemed clear that if compensatory help for children who need it is to have a lasting effect, it must start, not with a preschool program at age three or four, but in the home-preferably at birth, or before.

Dr. Sperber and his Brookline staff together with Dr. White and Harvard colleagues discussed specific plans for a school-based early education program. Since the young child's health is so vital to his learning, child health authorities from the Children's Hospital Medical Center were asked to join the planning group. The hospital encouraged by Dr. Julius Richmond, former national Director of Head Start and now Psychiatrist in Chief at Children's Hospital Medical Center, was concurrently becoming more involved with local community programs, and the matter at hand was certainly germane to that commitment.

After more than four years of discussions and planning, a pilot BEEP program emerged in 1973. Funding for the first two years of operation was obtained from the Robert Wood Johnson Foundation, and the Carnegie Corporation of New York. The program provides comprehensive diagnostic and education services for 300 families, beginning a few months before the birth of the child and continuing to the start of school.

DIAGNOSTIC SERVICES

The aim of the diagnostic services is to ensure that no child progresses through the preschool years with an undetected educational or physical handicap. BEEP provides free health and developmental examinations at the neighborhood BEEP center, beginning two weeks after the child is born.

During the first two and one half years the exams are at frequent intervals: 2 weeks, 3½ months, 6½ months, 11½ months, 14½ months, 24 months and 30 months. These early exams go beyond the typical pediatric exam by including neurologic screening, vision and hearing screening, and assessment of the child's mental development as well. Anomalies that require further attention are referred for evaluation to the proper specialist.

An important aspect of the exams is to enable the parents to become well-informed about their child's unique pattern of development, to gain insight into the normal range of spurts and lags in various areas of development, and to reserve judgments based on inadequate information or one time diagnosis. Following each exam, BEEP staff members review the child's pattern of development with the parents and send a report to the family pediatrician or health center. The review is more detailed as the child grows older and includes comments in several areas: neurologic evaluation, auditory screening, vision screening, growth and motor development, language, visual-perceptual-motor development, cognition, and social-behavioral development.

EDUCATION SERVICES

The aim of the BEEP education program is to help each child experience the best possible beginning in life by providing resources for the parents in their role as teachers of the young child. In serving as resources for parents, the intention is to increase the parent understanding of child development and to focus on the design of home conditions that encourage the child's emerging abilities.

The educational philosophy underlying BEEP does not aim to accelerate or force children's development. Instead, it is oriented toward assisting the family in arranging, for each child, an environment rich in resources and in providing opportunities for him to exercise his natural talents.

The project operates three educational program levels which vary in cost and from which other communities can later choose the one most appropriate for their needs. The three programs vary in the amount of scheduled contact between BEEP staff and parents. One program involves frequently scheduled seminars and home visits, as often as every two or three weeks; a second program has less frequent seminars and home visits, about once every four or six weeks; the third program has no formally scheduled seminars or home visits.

In all levels, each family is assigned a teacher to be the liaison between BEEP and the family. This teacher must have an academic background in child development, and must also be a parent.

Aside from the program intensity variation, all three programs provide many basic education services to all participating families. These include:

- calling upon their specially assigned teacher for information or help
- dropping in at the center whenever they like, bringing their children who will be cared for by trained staff in a specially equipped playroom
- exploring the materials about early childhood that BEEP has gathered together in its "resource center"
- borrowing books, pamphlets, and toys
- viewing films and video tapes on child development topics and on other aspects of childhood
- attending special events such as films on home safety, workshops on toy making, and discussion on a variety of fields related to child development
- using BEEP's free transportation service to and from the center
- learning about other resources for young children that exist in the Boston area - recreational, educational, and medical

ISSUES IN THE DESIGN OF THE BEEP EVALUATION

An experiment can be defined as a formal, self-conscious, purposive system aimed at learning from experience. When well conducted, it is characterized by experimenter control, precise treatments, and valid results. It usually involves an exploration of the effects of one active independent variable, such as a treatment or experimental program, on one or more dependent measures. The researcher controls all other disturbances in the system either through physical means such as selection, or probabilistically through the principle of randomization.

The natural temptation is to attempt to apply this structured world of inquiry to the study of social experiments. Several issues emerge in this application.

These issues embrace an inherent conflict between the logic of social innovation and the demands of classical experimental designs. This conflict is exacerbated when one attempts to fit a polymorphous social innovations such as BEEP into the "square box" of a cookbook experimental design. The best principles of this tradition would suggest in settings such as ours:

(i) identification of all eligible families,

(ii) selection of a random sample from this population,

(iii) random assignment of families from this sample to one of the three program levels or a control group.

The essence of this design is the principle of randomization in its two dimensions:

1) random sampling (step ii) and

2) random assignment (step iii)

Random assignment safeguards the internal validity of the experiment. When properly applied, this should guarantee that the experiment generates valid results. Random selection from the target population is a key component in the external validity of the experiment. It provides for the generalizability of experimental results to this target population.

Because of the intransigency of many social settings, the experimenter often lacks the necessary control to implement fully such a design. Campbell and Stanley in 1963 responded cogently to this problem with the notion of the quasi-experiment. They introduced 8 different designs and developed basic criteria for evaluating the usefulness of these designs in a specific research setting. This work evolved out of the classical tradition, and we are exhorted to maintain the trappings of that tradition wherever possible. While these developments have proved useful in small scale studies, such as curriculum evaluation in an individual school system, they have not been as helpful in the design of large-scale social experiments (Cohen, 1973).

The crux of the problem is that the criteria of good experimentation are often in conflict with the criteria of good service. This produces a tension between the notions of providing social services and learning from a structured experience. For example, social innovations such as BEEP require an intense personal commitment of both staff and participating families. This sense of commitment, however, is in conflict with notion of research objectivity. What constitutes a very desirable program objective, is also a serious threat to the generalizability of the experimental results.

In designing the BEEP study, we had to tussle with the very special problem of the randomly assigned control group. The principle of equal opportunity is fundamental to our society and in particular to our educational systems. In operation, this suggests that service programs should be universally and equally provided for all with maximum services perhaps going towards those with maximum need. The principle of random assignment is in conflict with this general ethic. Further, there are inevitably political and social problems in any attempt to implement random assignment. David (1974) recounts quite vividly the difficulties she encountered in her attempt to use random assignment in studying the effects of summer Follow Through programs. All of this adds up to the fact that random assignment is very difficult to employ in social settings. One can even go a step further, however, and raise the question of desirability. Even if random assignment is possible, will it achieve the intended goals?

When families agree to participate in a program such as BEEP, this is clearly an expression, on their part, of a perceived need for services, and not an expression of a desire to be part of some experiental learning process. They may accept the latter, but only as the necessary price of admission for the former. Thus, being assigned to a control group, even if by a truly random process, will not fulfill

their perceived needs. Inevitably, people will turn to other sources if available to fulfill these needs. Our control group may be someone else's treatment group. David (1974) presents a clear example of this problem too.

The problem is further compounded when the nature of the innovation involves societal change, as well as change in individual children. In statistical terms their are two units of analysis: the individual child and the community. Even if we can isolate the control group from individual services, we cannot isolate them from societal change.

In the BEEP setting our diagnostic services are technologically bound - they consist of highly advanced assessment procedures administered by trained professionals. They are discretely packaged, and we can direct the delivery of these services quite accurately. Our educational services, however, involve ideas and concepts. They are diffuse and their dissemination is impossible to control. Thus, families who are not in BEEP may still be affected by it either through broad societal changes, or a changing knowledge structure. The randomly assigned control group is a less than totally adequate approach in the face of these conditions. This approach is even less desirable when we confront the social, political, and ethical realities. It is a very difficult strategy to sell, and even if people buy it, it may not deliver the goods.

One alternative would be to employ a non-equivalent control group design (Campbell and Stanley, 1963) where the control group is a sample of children in another community hopefully similar to the experimental community. This approach might be useful with a large scale experiment consisting of multiple sites, if the experimental and control sites are well matched and if the children are also well matched across paired sites.

This is in fact the basic design of the Planned Variation Head Start experiment and the Follow Through study. Even in this setting, however the difficulties encountered are a sobering experience, and quickly temper ones enthusiasm for this approach (Smith, 1973; Weisberg, 1973). In more modest settings such as BEEP where there is only one experimental site, no approach to selecting a comparison community can provide adequate safeguards against possible selection biases.

This leaves us with only one viable alternative, an historical approach. For a short-term experiment, the historical approach uses the pre-existing condition in the experimental subjects themselves as the standard of comparison. For a longer term experiment, the comparison data would be the outcomes recorded on previous cohorts of children from the same community, but who were either ineligible or not exposed to the program. We have adopted this latter approach as the comparison baseline in the evaluation design for BEEP. For our first two evaluations, when the child is 14½ months and 30 months of age, we will gather evaluative baseline data on children born in 1972 in Brookline. These children were ineligible for BEEP because they were born prior to the initiation of the project. For our "in school" assessment, we will gather data on children born from 1967 to 1972 as they enter and progress through the first two grades of the Brookline Schools. These data together with comprehensive background data will form a time series which should provide an extensive base for estimating very precisely the long term effect of BEEP programs on child functioning in school. This approach provides estimates of the value-added by BEEP programs over and above the pre-existing condition. The full details of this data collection are presented in Figure 1.

There is still a problem in that BEEP families constitute a volunteer sample from the 1973 and 1974 cohorts. Thus, in theory, our sample may differ in innumerable ways on relevant background characteristics and attitudes from the population which constitutes our data baseline. We are, however, enrolling approximately 50% of all eligible families, and our early demographic analyses suggest that we have a representative cross-section of families from the target areas. What differences might exist should not be large in magnitude, and with our extensive data base we should be able to make adequate statistical adjustments. Further, we can exercise some control in the selection of subjects to form the data baseline. In particular, we can employ selection mechanisms which have a maximum likelihood of generating a baseline sample with the volunteer characteristics of the BEEP sample, while still controlling for measurable background variables.

Up until this point our primary concern has been the lack of experimenter control and the havoc it can create in an attempt to generate valid results. Having settled upon a design capable of producing valid results, we next consider the important issue of the generalizability of these results. This problem has several faces. First, there is the issue of voluntarism. With the exception of formal schooling between the ages of 6 and 16, participation in almost all social programs, stardard or innovative, is voluntary. People cannot be forced to participate. The notion of random selection from the target population while statistically elegant is socially abhorrent. Our only alternative is to redefine our target population to include only those families who might volunteer for a program such as BEEP. This definition is somewhat unsatisfying because it defines a population whose frame is subject to change as notions of acceptability change.

Second, as suggested earlier, complex social programs such as BEEP are organizational as well as technological innovations. The BEEP child is impacted not only by services directed specifically toward him and his family, but also by the amorphous changes occurring in the Brookline social network. At work here are delicate and complex interactions between an existing structure, its personalities, values, and traditions. The results of these interactions are contingent upon the specific nature of this structure and they are unique to it. When a program such as BEEP is attempted in only one setting, the notion of the generalizability of these societal changes is greatly restricted. Even if BEEP were conducted in 30 sites, although we might then have a better grasp of the possible range of outcomes and forces encountered, we would still be unable to make explicit statements about expected societal changes in another setting. The logic of these processes is far more complex than the logic of sampling; the generalizability of societal change based on the BEEP experience is fraught with errors.

Further, from a statistical point of view, we may find it useful to conceptualize the community social network factor in the perspective of a random effects rather than a fixed effects model. In this framework, the emphasis is clearly upon the variability in outcomes. The search for a precise estimate of central tendency, the generalizable effect, is not relevant.

Third, the notion of generalizing valid results is contingent upon the notion of a precise treatment. Unfortunately, in most social innovations, the program is not well-defined (Cohen, 1973). This lack of clarity is often due to the weak and uncertain theoretical foundation from which the innovation springs. Further, even if the theoretical underpinnings are sound, the program actually designed is the result of a complex system of adjustment among conflicting values.

It can often mean different things to different people, and rarely turns out as originally conceived. In addition, such innovations when first instituted may be very responsive to the social setting, and as a result often undergo major changes during implementation. Thus, it is important to analyze the stage of development of a particular innovation. If one has a field-tested well developed curriculum, the notion of a randomized trial may be more applicable. The desired outcome is clearly valid, generalizable results about the effects of the program on children. When we are considering, however, an innovation in its first stages of development, we must be equally concerned about the changing structure of the innovation itself, and its internal dynamic processes. How a complex program of services such as BEEP operates in fact, not just on paper, become a primary research objective.

This point of view has had a profound effect upon our project. Although BEEP was originally cast in a classical experimental framework (i.e., experimenter control, precise treatment, and valid generalizable results), we have found this framework inadequate. Although our diagnostic services are well defined and the desired outcomes are easily specified and measured, our educational services do not have these properties. The education programs are multi-faceted with initial emphasis on parent education. The theoretical basis of these services is drawn from the early childhood research of Burton L. White (1973) and others. This is, however, the first time that a comprehensive program of services of this type has ever been assembled in a public school setting. While these services are based on a theory about change, there is no evidence to suggest that they can in fact produce change, or that they are even responsive to perceived community needs. Further, our "treatment" is not uniform across parents, but rather is individualized for each family through a complicated network of interactions between our best pro-

fessional judgement of parents' needs and their own sense of personal needs. We offer many services, but the degree to which they are utilized and how they are utilized become important questions. If we cannot specify what is really going on within the program, then questions about program impact and its generalizabilty are without meaning. The importance of examining questions about the program content and its relevance as well as how parents perceive and respond to it, cannot be overstated.

With an innovation such as BEEP we are all novices. Thus, an introspective analysis constitutes perhaps the most useful and generalizable information that we can provide as we work through for the first time this complex innovation called BEEP. We can feedback this information to improve our own programs, and it should prove invaluable to others who might seek to initiate similar programs.

We have developed several research strategies to gather this process input:

1) unobtrusive data - frequency counts on parent utilization of various BEEP resources.

2) informal feedback by parents through our operational staff.

3) formal feedback by parents through:
   a) suggestion box
   b) parent group sessions
   c) individual parent interviews by an objective interviewer.

4) a project historian to document the developments in the entire program, and a teacher historian to focus specifically on changes within the education programs.

Each of these approaches has its own strengths and weaknesses, but taken together they should provide a well focused picture of the internal dynamics of BEEP.

In summary, we have come a long way from the classical notions of experimental design and generalizability of results. Our approach can be somewhat unsettling in that it lacks the definition and elegance of this experimental tradition. One might argue that we have made parents as well as their children the subject of this study.

This is true, but we too, the educators, pediatricians, and researchers are also subjects. It is a vital step to admit our own limitations and lack of answers. The development of effective child welfare programs is a far more complex task than the wisdom of the sixties would have suggested. With a firm grasp of the results of these past efforts, with modest expectations, and a sense of humility, however, we can move forward.

Figure 1: OVERVIEW OF BEEP DATA COLLECTION

Major Evaluations

Age Cohort	14½ Mo.	30 Mo.	School Entry	End of 2nd Grade
1974 BEEP	P	P	P	P
1974 NON BEEP[1]			C	C
1973 BEEP	P	P	P	P
1973 NON BEEP[2]			C	C
1972 COHORT	C	C	C	C
1971 COHORT			C	C
1970 COHORT			C	C
1969 COHORT			C	C
1968 COHORT				C
1967 COHORT				C

P = data gathered on program children
C = " " " comparison children

We have selected 14½ months, 30 months, entry into school, and end of the second grade as times points for major evaluations in which we compare BEEP children to comparison children. We are considering the desirability of adding another evaluation point between our 30 month and entry into school evaluations. Comparison data for this evaluation would probably be drawn from the 1972 - 1974 cohorts.

---

[1] The 1974 Non-BEEP group will consist of the following: children born after the recruitment deadline (tentatively Sept. 1, 1974), children of families not interested in BEEP, and children not enrolled in BEEP before the enrollment deadline (when the child is 2 weeks of age).

[2] The 1973 Non-BEEP group consists of the following: children born prior to the start of recruitment (March 1, 1973), children of families not interested in BEEP, and children not enrolled in BEEP before the deadline (when the child is 2 weeks of age).

Reference

Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research. In N.L. Gage (Ed.), Handbook of research on teaching. Skokie, Ill.: Rand McNally, 1963.

Cohen, D. Social experiments with schools: What has been learned? Paper presented at the Brookings Institute conference on social experimentation, Washington, 1973.

David, J. Final report on the summer effects study. Cambridge, Massachusetts: Huron Institute, 1973.

Smith, M.S. Some short term effects of Project Head Start: a preliminary report on the second year of planned variation: 1970 - 1971. Cambridge, Massachusetts: Huron Institute, 1973.

Weisberg, H.I. Short term cognitive effects of Head Start programs: a report of the third year of planned variation - 1971 - 1972. Cambridge, Massachusetts: Huron Institute, 1973.

White, B.L. & Watts, J.C. Experience and Envirnoment Major Influences on the Development of the Young Child. Volume 1. Englewood Cliffs, New Jersey; Prentice-Hall, 1973.